

## **GENERAL INFORMATION ON THE TRUMAI DIGITAL ARCHIVE**

### **HOW THE ARCHIVE WAS BUILT**

This digital archive was organized with recordings and materials obtained by three researchers who have worked directly with the Trumai people:

- Aurore Monod Becquelin, who started conducting her research on Trumai in 1966.
- Raquel Guirardello-Damian, who has been working with the group since 1989.
- Emmanuel de Vienne, who has been working with the Trumai since 2002.

The Trumai speakers themselves have contributed to the archive, either participating in the recordings or producing written texts and drawings. They also helped in the transcription of texts and translation into Portuguese (the national language of Brazil). The researchers further contributed with translation into French or English, and glossing and annotation of the texts.

With regard to the glossing, there are variations in the way certain morphemes are analyzed and glossed, given that the three researchers have different opinions about particular cases. However, despite these differences, the user of the archive will be able to understand the text and follow the glossing presented for it.

Next, we have information about the types of texts and data included in the Trumai corpus.

## TEXT TYPES

In the Trumai archive, there are texts of various genres: personal narratives, myths, descriptions, etc. Some of the texts have a recording (video or audio), some do not. The quality of the recording varies, because some texts were registered using an analog audio recorder, while others were registered using digital technology. The texts without a recording consist of materials written for educational purposes; some of them are actually short, but they are included in the archive because they are real instances of the use of the language. Some written texts have illustrations.

For the organization of the archive, the texts were divided into two main types:

(i) Monological: texts in which the main speaker talks continuously, without being interrupted by his/her audience. Sometimes a person of the audience may show signals that s/he is following the story (for example, by saying things like: "oh, really?" or "I see"), but the signal does not really interrupt the flow of the text.

(ii) Dialogical: texts in which two or more speakers interact. There are various turns, that is, speakers alternate turns (sometimes one speaker talks, sometimes the other one holds the talking).

It is important to note here that during the narration of myths, it may happen sometimes that a person of the audience poses a question to the narrator, and the monologue turns into a dialogue for a while, with interaction and exchange of turns. Thus, the division among the various kinds of texts is not so strict, but rather used for practical reasons (it is just for the question of how to organize the annotation. The annotation of a session with one main speaker is not so complicated; a session with two or more speakers interacting requires a little different organization). It is also the case that some Trumai myths may contain "non-linguistic" materials, such as songs (for example, when a character of the myth starts singing).

## **ELICITATION DATA**

The data from elicitation can be of various types:

(i) word lists: lexical items organized by semantic field. For example: names of plants; names of animals; terms for colors; kinship terms (mother, father, etc); verbs of motion; verbs of perception; verbs of transference, etc

(ii) sentences: these are sentences not from texts, but rather from questions made to the speaker during an elicitation session. The researcher makes the question in one language (Portuguese), and the consultant tries to give the requested data in Trumai.

(iii) stimuli: the elicitation is conducted with the help of visual stimuli, such as a video animation, or a series of pictures especially prepared for the elicitation.

(iv) field notes: it can be various kinds of information requested to the consultants (e.g. checking if some constructions are allowed in the language; checking a hypothesis). It involves longer answers from the consultant.

The data from elicitation in the Trumai archive is organized into two main subdivisions:

(i) data that was audio recorded; (ii) data that has no audio recording (only written register).

## SONGS

The Trumai people have various kinds of traditional songs. Some of them have lyrics that are in the Trumai language, such as the songs of the *Yepema'e* festivity. However, there are also songs with lyrics that are not in Trumai; for instance, some of the songs of the *Yamurikuma* ceremony, whose lyrics do not seem to be in any identifiable language (the lyrics of these songs have a special nature). The *Yamurikuma* songs were not created by the Trumai people; they learned them from other Upper Xingu tribes and incorporated them in their own tradition. In the corpus, there are other cases of songs such as these.

Given this scenario, it seems to be more appropriate not to include songs among the linguistic materials of the language (i.e., the texts, which are in Trumai), but rather to have a separate category for them. In this way, we can have the whole collection of traditional songs of the group presented together.

## STUDIES

The archive contains some studies about the Trumai language and culture. These studies are focused on aspects of the language (genetic affiliations, sound system, grammar) and aspects of the people (ethnographic information, sociocultural description, historical studies, etc). There are also materials of comparative studies between Trumai and other Xinguan languages, such as Kuikuro and Aweti.

The main goal of this section is to bring various kinds of information together, so that the user of the archive can have a better understanding about the Trumai group and its cultural aspects. Materials from various researchers will be included here (with acknowledgements).

## **METADATA**

Each session in the archive contains a metadescription with information about the content of the session, its participants, type of material, etc.

Given that all metadescriptions are open to the public, we do not identify the Trumai speakers by their full name. We use codes in order to keep their privacy. The archive organizers have further information about the code used for each speaker.